Genetic diversity in the Andes

GENETIC DIVERSITY IN THE ANDES: VARIATION WITHIN AND BETWEEN THE SOUTH AMERICAN SPECIES OF *OREOBOLUS* R. Br. (CYPERACEAE)

María Camila Gómez-Gutiérrez[1,2], R. Toby Pennington[1], Linda E. Neaves[1,3], Richard I. Milne[2], Santiago Madriñán[4] and James E. Richardson[1,5,†]

[1] Tropical Diversity Section, Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh, EH3 5LR, United Kingdom

[2] Institute of Molecular Plant Sciences, School of Biological Sciences, The University of Edinburgh, Daniel Rutherford Building, The King's Buildings, Edinburgh, EH9 3BF, United Kingdom.

[3] Australian Centre for Wildlife Genomics, Australian Museum Research Institute, Australian Museum, 1 William Street, Sydney 2010, Australia.

[4] Laboratorio de Botánica y Sistemática, Departamento de Ciencias Biológicas, Universidad de los Andes, Carrera 1 No. 18A – 10, Bogotá, Colombia.

[5] Programa de Biología, Universidad del Rosario, Carrera 26 No. 63B – 48, Bogotá, Colombia.

[†] Corresponding author, email: jamese.richardson@urosario.edu.co

ABSTRACT

This study examines genetic relationships among and within the South American species of *Oreobolus* that span the temperate and tropical Andes hotspots and represent a good case study to investigate diversification in the Páramo. A total of 197 individuals covering the distributional range of most of these species were sequenced for the nuclear ribosomal internal transcribed spacer (ITS) and 118 individuals for three chloroplast DNA regions (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L). Haplotype networks and measures of genetic diversity were calculated at different taxonomic and geographic levels. To test for possible geographic structure, a Spatial Analysis of Molecular Variance (SAMOVA) was undertaken and species relationships were recovered using a coalescent-based approach. Results indicate complex relationships among the five South American species of *Oreobolus*, which are likely to have been confounded by incomplete lineage sorting, though hybridization cannot be completely discarded as an influence on genetic patterns, particularly among the northern populations of *O. obtusangulus* and *O. cleefii*. We report a case of cryptic speciation in *O. obtusangulus* where northern and southern populations of morphologically similar individuals are genetically distinct in all analyses. At the population level, the genetic evidence is consistent with contraction and expansion of islands of Páramo vegetation during the climatic fluctuations of the Quaternary, highlighting the role of these processes in shaping modern diversity in that ecosystem.

KEYWORDS

Biogeography, Andes, species tree, lineage sorting, hybridization, Páramo.

ACKNOWLEDGEMENTS

This work was funded by a School of Biological Sciences Scholarship provided through The University of Edinburgh. We thank the herbaria at Aarhus University, (Denmark), Naturalis (The Netherlands) and Reading University (Great Britain) for making material available for DNA extraction. We also thank three anonymous reviewers for their valuable comments and James Nicholls from The University of Edinburgh for assistance with the *BEAST analysis.

INTRODUCTION

The Páramo is a putatively young ecosystem that appeared following the final uplift of the northern section of the Andes Mountain Range during the Pliocene, c. 5 million years ago – Ma (van der Hammen 1974; van der Hammen and Cleef 1986; Hooghiemstra et al. 2006; Graham 2009). It occupies an area of 37500 km^2 and is distributed in a series of sky islands with c. 4000 plant species of which 60% are endemic (Luteyn 1999; Buytaert et al. 2010). It has been proposed that the glacial-interglacial cycles of the Quaternary may have played an important role in shaping Páramo plant populations (van der Hammen 1974; Simpson 1975). The continuous contraction and expansion of altitudinal vegetation belts may have promoted the

contact of Páramo islands during glacial periods, enabling the migration and exchange of otherwise isolated taxa (van der Hammen and Cleef 1986). Conversely, during interglacial periods, Páramo islands may have been isolated, promoting speciation (van der Hammen and Cleef 1986). Furthermore, previous studies have demonstrated that Páramo lineages have significantly higher speciation rates than any other biodiversity hotspot on Earth and that many speciation events occurred during the Pleistocene (Madriñán et al. 2013). Recent divergence times among Páramo plant lineages might have implications, both at the phenotypic and genotypic level, because morphological diversity and differentiation may not reflect complete genetic divergence between and within closely related taxa (Schaal et al. 1998).

The five South American species of the schoenoid sedge *Oreobolus* R. Br. (*O. cleefii* L.E. Mora, *O. ecuadorensis* T. Koyama, *O. goeppingeri* Suess., *O. obtusangulus* Gaudich. and *O. venezuelensis* Steyerm.) are an ideal model system to investigate how recent climatic and/or geological events may have shaped extant populations in the Páramo. Previous studies have supported the monophyly of the South American clade of *Oreobolus* and dated its divergence to c. 5 Ma, coinciding with the appearance of the Páramo ecosystem (Chacón et al. 2006). The South American clade of *Oreobolus* is therefore a good exemplar to study Páramo biogeography, including investigating the likely effects of recent climatic events (i.e. glacial cycles of the Quaternary) on the population structure of its species.

A handful of genetic studies for similar high-altitude tropical ecosystems in Africa have been published in recent years (Kebede et al. 2007; Assefa et al. 2007; Gizaw et al. 2013; Kadu et al. 2013; Wondimu et al. 2013). However, such studies are almost non-existent for the Páramo flora (Vásquez et al. 2016; Kolář et al. 2016). The aims

of this study are to estimate the species phylogeny of the South American species of *Oreobolus* and their timing of diversification, to assess genetic structure at the inter- and intra-specific level and to interpret these in the light of Quaternary glacial-interglacial cycles.

METHODS

Study species and sampling

The species concepts for *Oreobolus* that we use here follow the monograph of Seberg (1988) for *O. ecuadorensis*, *O. goeppingeri*, *O. obtusangulus* and *O. venezuelensis*, and of Mora-Osejo (1987) for *O. cleefii*. These species, with the exception of *O. obtusangulus* subsp. *obtusangulus,* are restricted to wet, temperate-like environments in the northern section of the Tropical Andes and in the Talamanca Cordillera in southern Central America, and are found only in the high-altitude Páramo ecosystem (Seberg 1988; Chacón et al. 2006). *Oreobolus cleefii* is restricted to the Eastern Cordillera and the southern Andean region of Colombia. *Oreobolus ecuadorensis* is found in southern Colombia, Ecuador and northern Peru. *Oreobolus goeppingeri* is distributed in the Talamanca Cordillera in southern Central America, Colombia and Ecuador. *Oreobolus obtusangulus* has two subspecies with a disjunct distribution: subsp. *unispicus* is distributed in Colombia, Ecuador and northern Peru while subsp. *obtusangulus* occupies the subantarctic region of Chile, Argentina and the Falkland Islands. Finally, *O. venezuelensis* occupies all Páramo regions (Talamanca Cordillera, Venezuela, Colombia, Ecuador and northern Peru).

The distributions of all *Oreobolus* Páramo species overlap with those of at least one other congeneric species (Fig. 1). All Páramo species are found between 3000 and 4300 m a.s.l. while in the subantarctic regions, the altitude at which *O. obtusangulus* is found decreases with increasing latitude, from 2400 m a.s.l. to sea level (Seberg 1988). The five South American species are clearly differentiated in terms of morphology and, in common with most Cyperaceae, *Oreobolus* is both wind pollinated and dispersed (Seberg 1988). Little is known about ploidy levels and chromosome numbers in *Oreobolus*, with the only chromosome count for *O. obtusangulus* ssp. *obtusangulus* (2n = 48; Moore (1967)).

The five South American species of *Oreobolus* (*O. cleefii*, *O. ecuadorensis*, *O. goeppingeri*, *O. obtusangulus* and *O. venezuelensis*) were sampled extensively across their entire distribution range (Fig. 1). A total of 269 samples from 32 sampling localities were obtained from both field collections (10 sampling localities) and herbarium material (22 sampling localities) (Fig. 1 and Supp. Table 1). From each of the ten field sampling localities, all within Colombia, two to ten fresh leaf samples per species were collected, and their location was recorded using a handheld GPS (Fig. 1, sampling localities 2 – 11). For sampling localities in Costa Rica, Ecuador, Peru, Chile and Argentina (Fig. 1, sampling localities 1 and 12 – 32), herbarium material was acquired from the Utrecht (U) and Leiden University (L) branches of the National Herbarium of the Netherlands, Aarhus University Herbarium (AAU) and the University of Reading Herbarium (RNG). For herbarium specimens, between one and ten individuals per species were sampled from each sampling locality. Coordinates were recorded from the herbarium specimens and checked for accuracy using the NGA GEOnet Names Server (GNS) (http://geonames.nga.mil). Sampling

localities are numbered 1 to 32 in a north to south direction. Sampling localities 1 to 23 will be referred to as northern Andes – NA (Costa Rica, Colombia, Ecuador and Peru) and 24 to 32 as southern Andes – SA (Chile and Argentina). Previously published sequence data for *O. cleefii*, *O. goeppingeri* and *O. venezuelensis* (Chacón et al. 2006) were also incorporated and assigned to their corresponding sampling locality. Supplementary Table 2 presents the complete list of samples used in this study together with their GenBank numbers.

DNA extraction, amplification and sequencing

Both silica-dried fresh leaf samples and herbarium material were pulverised using a Mixer Mill (Retsch, Haan, Germany). Total genomic DNA from herbarium material was isolated following the CTAB method of Doyle and Doyle (1990) and from silica-dried samples with the DNeasy® Plant Mini Kit (QIAGEN, Manchester, UK) following the manufacturer's protocol. The chloroplast region *trn*L-F was amplified and sequenced using primers *trn*Lc and *trn*Lf for silica-dried material, and in combination with internal primers *trn*Ld and *trn*Le for herbarium material (Taberlet et al. 1991). For silica-dried material, the ITS region was amplified and sequenced with external primers ITS5P and ITS8P (Möller and Cronk 1997). For herbarium material, owing to the increased likelihood of the DNA being degraded, amplification and sequencing were performed using external primers ITS5P and ITS8P in combination with internal primers ITS2P and ITS3P (Möller and Cronk 1997), in order to amplify the shorter ITS1 and ITS2 regions in separate reactions. The chloroplast regions *trn*H-*psb*A and *rpl*32-*trn*L were amplified and sequenced using primer pairs *trn*H^{GUG} (Tate and Simpson 2003)/*psb*A (Sang et al. 1997) and

*trn*L$^{(UAG)}$/*rpl*32-F (Shaw et al. 2007), respectively. For all reactions, 20 µl PCR reactions used the following proportions: 1 µl of unquantified DNA, 1x Buffer (Bioline, London, UK), 1mM dNTPs, 1.5 mM $MgCl_2$ (Bioline, London, UK), 0.75 µM of each forward and reverse primer, 4µl of combinatorial enhancer solution (CES) and 0.05 U of *Taq* polymerase (Bioline, London, UK). The amplification cycle for all chloroplast regions (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L) consisted of 2 min at 94 ºC, followed by 30 cycles of 1 min at 94 ºC, 1 min at 52 ºC and 1 min at 72 ºC, finalising with 7 min at 72 ºC. For ITS, the amplification cycle consisted of 3 min at 94 ºC, followed by 30 cycles of 1 min at 94 ºC, 1 min at 55 ºC and 90 sec at 72 ºC, finalising with 5 min at 72 ºC. PCR products were purified with 2 µl of ExoSAP-IT® (USB Corporation, High Wycombe, UK) for 5 µl of product. Sequencing reactions for each primer used the BigDye® Terminator v3.1 chemistry (Applied Biosystems™, Paisley, UK) and the manufacturer's protocol. Sequencing was performed at the Edinburgh Genomics facility of the University of Edinburgh. No double peaks were observed in the chromatograms of the ITS region and therefore it was not necessary to clone.

Matrix assembly and sequence alignment

Contigs of forward and reverse sequences were assembled in Sequencher version 5.2 (Gene Codes Corporation, Ann Arbor, Michigan, USA). 230 ITS sequences, 169 *trn*L-F sequences, 128 *trn*H-*psb*A sequences and 190 *rpl*32-*trn*L sequences were generated for this study (Supp. Table 2). The sequences were manually aligned using Mesquite v2.75 (Maddison and Maddison 2014). Supplementary Table 3 describes

number of individuals successfully sequenced per species per cluster/sampling locality.

Species phylogeny and timing of diversification

The multispecies coalescent model implemented in *BEAST 2 (Heled and Drummond 2012; Bouckaert et al. 2014) was used to estimate the phylogenetic relationships amongst the five South American species of *Oreobolus* as well as their divergence time. Only complete sequences were used for the species tree estimation (ITS, *trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L; Supp. Table 2). The analysis was run using bModelTest (Bouckaert and Drummond 2017) which is a model selection tool incorporated in BEAST 2 (Bouckaert et al. 2014) that uses a Bayesian framework (reversible jump MCMC) to select the most appropriate substitution model while simultaneously estimating the phylogeny. Phylogenetic reconstruction and divergence time estimations were performed using BEAST v2.4.5 (Bouckaert et al. 2014). The tree model was linked for the three plastid regions because cpDNA does not undergo recombination. The model of lineage-specific substitution rate variation was set as a strict clock model for each dataset. A *BEAST analysis requires each taxon to be associated with a species or taxonomic unit (Taxon Sets). These were defined following current taxonomy but with *O. obtusangulus* divided into northern and southern taxa (based upon results presented below). The diversification model for the species tree was set to a calibrated Yule model (Heled and Drummond 2012) with the population size model at its default setting. The root of the species tree was clock calibrated using a prior with a normal distribution defined by a mean (μ) of 4.76 Ma and a standard deviation (σ) of 1.2 Ma. The age and error range correspond

to those estimated for the crown node of the South American *Oreobolus* clade from a dated phylogeny of the Schoeneae tribe using one fossil and one secondary calibration (Gómez-Gutiérrez, 2016). A normal distribution was used on the root because it is the most suitable for secondary calibrations (Ho and Phillips 2009). This type of distribution allocates most of the probability density around the mean and allows for symmetrical decrease towards the tails accounting for age error (Ho and Phillips 2009). All other priors were left at their default settings.

Four independent MCMC runs of 250 million generations each were performed, sampling every 25000 generations. Runs were combined and 75% of the samples were discarded as burn-in. Adequate mixing and convergence were assessed using Tracer v1.6.0 (Rambaut et al. 2013). A maximum clade credibility tree (MCC) from the combined tree sets was annotated with common ancestor heights, 95% HPD node ages and posterior probability values (PP) on TreeAnnotator v2.1.2 (Rambaut and Drummond 2015).

Haplotype definition and networks

Haplotypes were identified independently for the nuclear ribosomal region (ITS) and the concatenated plastid regions (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L) in Microsoft Excel (Microsoft Corporation, Washington DC, USA) using the Chloroplast PCR-RFLP Excel macro (French 2003). For ITS, only samples successfully sequenced for the whole region were included (Supp. Table 2). Likewise, for the concatenated plastid regions, only samples successfully sequenced for all three regions were considered (Supp. Table 2). Informative insertion/deletion events (indels) were included in the analysis and coded as absent (0) or present (1) following the simple

indel coding method of Simmons and Ochotenera (2000). Poly-T and poly-A length polymorphisms, di-nucleotide repeats and ambiguously aligned regions were excluded from subsequent analyses for all regions. Haplotype connection lengths were calculated using Arlequin ver3.5 (Excoffier and Lischer 2010) and a minimum-spanning tree was produced in Hapstar v0.5 (Teacher and Griffiths 2011).

NeighborNet networks – NN (Bryant and Moulton 2004) were also constructed for both nuclear and concatenated plastid haplotypes using Splitstree 4 (Huson and Bryant 2006). This method allows representation of conflicting signals in the data, which might be due to incomplete lineage sorting or reticulate evolution (Bryant and Moulton 2004; Huson and Bryant 2006). In the resulting network, conflicts are represented by parallel edges connecting taxa. The NN networks used uncorrected-p distances, which calculate the number of changes between each pair of haplotypes.

Genetic diversity and structure

Sampling localities were combined into clusters to increase the likelihood of detecting phylogeographic signal (Fig. 1, Supp. Table 1). Clusters were defined regardless of species classification, an approach justified by Gómez-Gutiérrez (2016; see also results below) who showed poor phylogenetic resolution amongst the South American species of *Oreobolus*. Fourteen clusters (A – N) were defined according to geographic distance and ensuring the absence of any significant geographic barrier between sampling localities within each cluster such as deep inter-Andean valleys.

Haplotype (h) and nucleotide (π) diversities were calculated independently for each cluster and each species in Arlequin ver3.5 (Excoffier and Lischer 2010).

Additionally, haplotype richness (hr) was estimated for each species using HIERFSTAT (Goudet 2005) in R version 3.2.3 (R Core Team 2015). This package uses a rarefaction procedure set to 100 runs to correct for bias due to unequal sample sizes. ITS sample size was standardised to 15 individuals while cpDNA sample size was standardised to nine. Additionally, F_{ST} values between cluster pairs and species pairs were calculated independently for ITS and the concatenated plastid regions using Arlequin ver3.5 (Excoffier and Lischer 2010). NN networks for both nuclear and concatenated plastid regions were constructed from the calculated F_{ST} values. For the cluster pairs, clusters A, K and N were excluded from the analysis due to their low sample sizes ($N \leq 2$). In the case of the species pairs, calculations were first undertaken considering *O. obtusangulus* as one species and then with the northern and southern populations considered as two different species.

To analyse the geographical structure of genetic variation, a spatial analysis of molecular variance (SAMOVA) was performed independently for the nuclear and concatenated plastid datasets (Dupanloup et al. 2002). SAMOVA identifies groups of populations/clusters that are geographically homogeneous as well as maximising genetic differentiation amongst them (Dupanloup et al. 2002). One hundred annealing simulations were undertaken for each possible number of groups (ITS, K = 2 – 13; cpDNA, K = 2 – 12). The minimum number of groups (K) was chosen that maximised the genetic differentiation amongst them (F_{CT}). Subsequently, haplotype (h) and nucleotide (π) diversities were calculated for the resulting SAMOVA groups in Arlequin ver3.5 (Excoffier and Lischer 2010). Likewise, haplotype richness (hr) was estimated for each group using HIERFSTAT (Goudet 2005) in R version 3.2.3 (R Core Team 2015). Similarly, to test if the phylogeographic structure had a

phylogenetic component, two measures of genetic differentiation amongst clusters were estimated using PERMUTCPSSR 2.0 (Pons and Petit 1996; Burban et al. 1999). A distance matrix was calculated based on the number of mutational steps between haplotypes (N_{ST}) and on haplotype frequencies (G_{ST}). Ten thousand permutations were performed to assess if N_{ST} was significantly higher than G_{ST}.

Additionally, variation in genetic structure was further examined for 1) all species, 2) all clusters, 3) northern Andes clusters only, 4) clusters grouped by region (northern Andes, southern Andes) and 5) SAMOVA groups using an analysis of molecular variance (AMOVA) in Arlequin ver3.5 (Excoffier and Lischer 2010).

RESULTS

Species phylogeny and timing of diversification

The MCC tree for the combined tree sets (Fig. 2) shows *O. cleefii*, *O. ecuadorensis*, *O. goeppingeri* and *O. venezuelensis* are recovered as monophyletic. The results support the genetic differentiation between *O. obtusangulus* from the northern Andes region (NA; Fig. 2) and *O. obtusangulus* from the southern Andes region (SA; Fig. 2). *Oreobolus obtusangulus* (SA) is sister to all remaining species. In the northern Andean clade (NAC; PP=100%), *O. ecuadorensis*, *O. cleefii* and *O. obtusangulus* (NA) form a clade (PP=75%) sister to another clade composed of *O. goeppingeri* and *O. venezuelensis* (PP=71%). *Oreobolus cleefii* and *O. obtusangulus* (NA) are recovered as sister species (PP=87%). South American *Oreobolus* diverged c. 4.39 Ma (95% HPD [1.96 – 6.97] Ma) during the Pliocene (Fig. 2). Subsequently, the

NAC diversified into five species c. 0.44 Ma (95% HPD 0.11 – 0.81] Ma) during the Pleistocene (Fig. 2).

Haplotype definition and networks

Nuclear ribosomal DNA

A total of 197 individuals from 14 clusters (A – N) were scored for ITS haplotypes, including individuals for all five species across their entire distribution range (Supp. Table 3). After exclusion of poly-T and poly-A length polymorphisms, di-nucleotide repeats and ambiguously aligned regions, 523 bp of aligned sequences remained. Thirty-nine polymorphic sites comprising 38 nucleotide substitutions and one indel defined thirty haplotypes. Of these, 22 (73.3%) were species-specific while eight (26.7%) were shared among species (Fig. 3, Supp. Table 4 and Supp. Fig. 1). There was no clear clustering according to current taxonomy evident in either the NeighborNet network (NN) (Fig. 3) or the minimum-spanning tree (MST) (Supp. Fig. 1), for example *O. obtusangulus* is not resolved in one group.

At a continental scale, haplotypes were geographically restricted with no shared haplotypes between the NA region and the SA (Fig. 3, Supp. Table 4 and Supp. Fig. 1). This geographic structure was evident in both the minimum-spanning tree (Supp. Fig. 1) and the NN network (Fig. 3). Within the NA sampling localities, patterns were more complicated. There are eight shared haplotypes evident in the MST (Supp. Fig. 1) and many edges in the NN Network (Fig. 3). Of the eight shared haplotypes, seven occur in *O. obtusangulus*. Furthermore Hn9, a haplotype shared between *O. goeppingeri* and *O. obtusangulus*, is located in the middle of the MST connecting the SA and NA haplotypes (Supp. Fig. 1). When not considering shared

haplotypes, Hn12 and Hn14 found in *O. goeppingeri* are closer to those found in other species than they are to other haplotypes of the same species as are Hn28 and Hn30 in *O. venezuelensis*.

Plastid DNA

A total of 118 individuals from 13 clusters (B – N) were successfully sequenced for all three plastid markers (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L), including individuals from all five species across most of their distribution range (Supp. Table 3). A concatenated matrix of 2465 bp of aligned sequences (*trn*L-F, 1040 bp; *trn*H-*psb*A, 676 bp; *rpl*32-*trn*L, 749 bp) resulted after the exclusion of poly-T and poly-A length polymorphisms, di-nucleotide repeats and ambiguously aligned regions. Forty haplotypes were identified based on 141 polymorphic sites (*trn*L-F, 53; *trn*H-*psb*A, 14; *rpl*32-*trn*L, 74) including 112 nucleotide substitutions and 28 indels. Thirty-four haplotypes (85%) were species-specific while six (15%) were shared among species (Fig. 4, Supp. Table 5 and Supp. Fig. 2). When only considering species-specific haplotypes, both the MST and NN network showed some degree of clustering according to taxonomy for three of the species, namely *O. ecuadorensis*, *O. goeppingeri* and *O. venezuelensis* (Fig. 4 and Supp. Fig. 2).

As for ITS, there were no shared haplotypes between the NA and the SA regions (Fig. 4, Supp. Table 5 and Supp. Fig. 2). This geographic structure was evident in both the MST and the NN network (Fig. 4 and Supp. Fig. 2). There was low support for groupings in the cpDNA network in the relationships amongst NA groups compounded by the large number of possible unsampled haplotypes. The results of

the cpDNA analysis were similar to those of ITS in showing a large number of edges and of shared haplotypes.

Genetic diversity and structure

Species genetic structure

Molecular diversity indices for ITS and cpDNA for the five *Oreobolus* species, including the two *O. obtusangulus* groups (NA and SA), are shown in Table 1. Haplotype and nucleotide diversity was lowest in *O. ecuadorensis* and highest in *O. obtusangulus* (Table 1). Similarly, haplotypic richness was lowest in *O. ecuadorensis* and highest in *O. obtusangulus*. However, the high values in *O. obtusangulus* were reduced when considering SA and NA populations of *O. obtusangulus* as separate species (see Table 1).

Pairwise F_{ST} values between all species pairs were significant for ITS and cpDNA (ITS: $p < 0.001$; cpDNA: $p < 0.05$), with the exception of *O. cleefii* and *O. obtusangulus* (NA) for cpDNA ($F_{ST} = -0.020$) (Table 2, Supp. Figs 3 – 4). *Oreobolus ecuadorensis* is consistently differentiated from the other species in both ITS and cpDNA (Table 2, Supp. Figs 3 – 4). The NN, based on F_{ST} values showed that when considering *O. obtusangulus* as one species, it is reconstructed in the middle of the network and its placement is poorly resolved in both ITS and cpDNA NN networks (Supp. Figs 3a – 4a). In contrast, when considering northern and southern groups separately, *O. obtusangulus* (SA) is clearly different from other *Oreobolus* species, whereas *O. obtusangulus* (NA) has affinities with *O. cleefii*. The conflicting signal between the latter two species (i.e., multiple parallel edges) is evident in both cpDNA and ITS NN networks (Supp. Figs 3b – 4b). *Oreobolus*

goeppingeri and *O. venezuelensis* are well differentiated in cpDNA but not in ITS where they appeared in the centre of the networks with multiple connections to the other species (Supp. Figs 3 – 4).

Cluster genetic structure

The results of the AMOVA showed that although differentiation amongst species was significant (ITS, $F_{ST} = 0.30$, $p < 0.001$; cpDNA, $F_{ST} = 0.48$, $p < 0.001$), within species variation accounted for 70% for ITS and 52% for cpDNA (Table 3). Similarly, separation into geographic clusters only explained 43% (ITS) and 37% (cpDNA) of the variation.

The SAMOVA for both ITS and cpDNA indicated three groups (I – III; Supp. Table 8, Supp. Figs 1 – 2) as the number of genetic clusters (K) that maximised genetic differentiation amongst groups while minimising the number of single-cluster groups (ITS, $F_{CT} = 0.622$, $p < 0.001$; cpDNA, $F_{CT} = 0.426$, $p < 0.001$). For ITS, group I included all NA clusters (A – J) while groups II (K, L, N) and III (M) included the SA ones (Supp. Table 8, Supp. Fig. 1). For cpDNA, group I included all NA clusters plus the northernmost SA cluster (K), while groups II (L, N) and III (M) included the rest (Supp. Table 8, Supp. Fig. 2). SAMOVA groups explained slightly more of the genetic structure (ITS, $F_{CT} = 0.62$, $p < 0.001$; cpDNA, $F_{CT} = 0.43$, $p < 0.001$) than the NA versus SA continental divide (ITS, $F_{CT} = 0.60$, $p < 0.001$; cpDNA, $F_{CT} = 0.36$, $p < 0.001$) (Table 3). Molecular diversity indices calculated for the SAMOVA groups are presented in Table 4. Significant phylogeographic structure was indicated by the significantly higher values of N_{ST} (ITS, $N_{ST} = 0.605$; cpDNA, $N_{ST} = 0.406$) compared to G_{ST} (ITS, $G_{ST} = 0.262$; cpDNA, $G_{ST} = 0.156$; $p < 0.01$).

DISCUSSION

Timing of diversification

The dated species tree presented here (Fig. 2) indicates younger diversification dates than those presented by Chacón et al. (2006), which is expected because divergence dates estimated from a species tree will generally be younger than those estimated from a gene tree (Drummond and Bouckaert 2015). Our species phylogeny indicates that the most recent common ancestor of the South American *Oreobolus* diverged 4.39 Ma (95% HPD [1.96 – 6.97] Ma) during the late Miocene – early Pliocene. Subsequently, the northern Andean clade (NAC) appears to have diversified from 0.44 Ma (95% HPD [0.11 – 0.81] Ma). This indicates that the expansion and contraction of Páramo islands during the glacial cycles of the Quaternary may have played a role in diversification in the northern Andes (see last section of the discussion) (van der Hammen 1974; Simpson 1975; van der Hammen and Cleef 1986; Hooghiemstra and van der Hammen 2004).

Genetic diversity and structure

Our results reveal a complex evolutionary history for the five South American species of *Oreobolus*. Species relationships were difficult to estimate, indicating either interspecific gene flow and/or incomplete lineage sorting (Naciri and Linder 2015). Haplotype and nucleotide diversity were high for both ITS and cpDNA for all species except *O. ecuadorensis* (Table 1). Additionally, shared haplotypes were observed in both ITS (27%) and cpDNA (15%). This intricate history is also evident

in the MST and NN networks for both ITS and cpDNA (Figs. 3 – 4 and Supp. Figs. 1 – 2).

The high degree of complexity observed amongst these species contrasts with the morphological characters that distinguish them. Inconsistencies between morphological characteristics and genetic patterns can arise due to high levels of plasticity of morphological characters or parallel adaptations to local conditions resulting in the same morphology, which might be the case for *O. obtusangulus*. The data presented here indicate that the two subspecies of *O. obtusangulus* represent morphologically cryptic species. Britton et al. (2014) have described another example of cryptic speciation within the Schoeneae in the South African species *Tetraria triangularis*. These authors found at least three intraspecific lineages that qualified as cryptic species based on their genetic distinctiveness and subtle morphological differentiation. Furthermore, cryptic lineages have also been found in otherwise morphologically indistinguishable taxa within the Páramo genus *Loricaria* (Asteraceae) (Kolář et al. 2016).

Nonetheless, convergent morphological evolution does not appear to satisfactorily account for the genetic patterns observed in many South American species of *Oreobolus*, which may result from incomplete lineage sorting (ILS) and/or hybridization. Given the recent Pliocene diversification of both the northern and southern Andean clades of *Oreobolus* (Fig. 2), lineage sorting may not have been fully completed. Previous studies have indicated ILS in recently diverged groups, particularly when effective population sizes are large (Maddison and Knowles 2006; Jakob and Blattner 2006; Degnan and Rosenberg 2009; Cutter 2013). Furthermore, under a scenario of ILS, it is expected that different genes would have different

coalescence times. Haploid plastid genes have a lower effective population size than nuclear genes and thus would coalesce faster (Schaal and Olsen 2000; Naciri and Linder 2015). Faster coalescence would be translated into an increased correspondence between the genetic relationships recovered with plastid genes and currently recognised taxonomic species. Our results support this scenario because cpDNA better differentiates taxonomic species than ITS for *O. ecuadorensis*, *O. goeppingeri* and *O. venezuelensis* (Figs. 3 – 4 and Supp. Figs. 1 – 2).

However, species relationships may be obscured by ongoing gene flow as patterns of ILS are difficult to disentangle from those of historic hybridization. Two species pairs, *Oreobolus cleefii* and *O. obtusangulus* (NA), and *O. goeppingeri* and *O. venezuelensis*, show patterns indicative of ILS and/or hybridization. Firstly, *Oreobolus cleefii* and *O. obtusangulus* (NA) show contrasting patterns between nuclear (ITS) and cpDNA haplotypes (Figs. 3 – 4 and Supp. Figs. 1 – 2) possibly due to chloroplast capture and simultaneous nuclear introgression (Abbott et al. 2013). These closely related species naturally occur in sympatry in all of the sampled localities (Fig. 1) and show an overlap in morphological characters (Seberg 1988). In fact, morphological similarities previously lead Seberg (1988) to suggest that *O. cleefii* should be reduced to synonymy under *O. obtusangulus* subsp. *unispicus*, the northern Andean subspecies of *O. obtusangulus*. Secondly, the two most widespread species in the Páramo, *O. goeppingeri* and *O. venezuelensis*, also naturally occur in sympatry in all sampled localities (Fig. 1). These species also show complicated genetic patterns, combining high levels of diversity with shared haplotypes (Figs. 3 and 4) and conflicting phylogenetic relationships (Supp. Figs. 3 – 4 with other northern Andean species). A possible explanation is that the widespread nature of

these species provided greater opportunities for intra, and interspecific mixing compared with more range-restricted species, which exhibit a similar pattern of haplotype sharing, albeit on a smaller scale (Supp. Figs. 1 – 2).

Current gene flow would be expected to result in F1 hybrids that would exhibit heterozygosity in ITS, but this was not observed in any *Oreobolus* species, although such heterozygosity may no longer be evident in older hybrids. While the presence of later generations of hybrids or backcrosses cannot be excluded, the lack of heterozygosity in ITS and the presence of shared haplotypes recovered in multiple pairs of individuals from all species is more suggestive of a stochastic process likely related to lineage sorting. Therefore, although gene flow cannot be ruled out and may have a role in some situations (e.g. *Oreobolus cleefii* and *O. obtusangulus* see below), we suggest incomplete lineage sorting in a recently diversified group is also part of the explanation for the complex patterns observed in the South American species of *Oreobolus*. A recent phylogeographic study of the Australian alpine *Poa* (Poaceae) describes a similar pattern of problematic recovery of species relationships associated with a putatively young ecosystem and a Pleistocene radiation following long-distance dispersal to Australia (Griffin and Hoffmann 2014). This study also favoured ILS rather than ongoing gene flow as the likely process behind the observed pattern based on the widespread genetic similarity and recent divergence times.

The results of the AMOVAs revealed that neither clustering into currently defined taxonomic species (Mora-Osejo 1987; Seberg 1988) nor into our pre-defined geographic clusters (Fig. 1, Supp. Table 1) described the distribution of genetic diversity, only explaining 30% (ITS)/48% (cpDNA) and 43% (ITS)/37% (cpDNA),

respectively (Table 3). Rather, the SAMOVA suggested that an *a posteriori* geographic arrangement better explained genetic diversity (62% for ITS and 43% for cpDNA, Table 3). Thus, the observed patterns of genetic diversity are likely to be the result of complex interactions between some species over various geographic distances.

At a continental scale there is evidence of geographic structure in *Oreobolus* species, (Figs. 1 – 2, Supp. Figs. 1 – 2), suggested by a higher value of N_{ST} compared to G_{ST} ($p < 0.01$), indicating that haplotypes in the same cluster are on average more closely related than distinct haplotypes from different clusters. The clearest geographic break apparent in *Oreobolus* is between the northern Andes (NA) and southern Andes (SA). This pattern is evident in both chloroplast and nuclear regions, although the pattern is much stronger in ITS (Figs. 1 – 2, Supp. Figs. 1 – 2). The arid central Andes are likely to impose a barrier to dispersal and gene flow, but the position of the north-south break is unclear. SAMOVA groups clearly identify the NA/SA disjunction in ITS but not in the plastid region where cluster K is grouped with the northern Andean clusters (Supp. Table 8, Supp. Figs. 1 – 2). The latter is also evident in the cpDNA NN where the distance between haplotypes is shorter than in the NN for ITS (Figs. 1 – 2). The incongruence between ITS and plastid regions may suggest mixing between the SAC and NAC in cluster K, resulting from long distance dispersal events. Cluster K is separated from both NA clusters and other SA clusters by a substantial distance and possesses unique haplotypes at both ITS and plastid regions (Supp. Tables 4 – 5).

Additional structure is evident at regional scales within the NAC and appears to be associated with putative geographic barriers to gene flow. Pairwise F_{ST} values

calculated for ITS showed that clusters B and J were significantly differentiated from all other sites, regardless of the geographic distances (Fig. 5, Supp. Table 7). These two clusters are separated from all other NA clusters by inter-Andean valleys of seasonally dry tropical forest. Cluster B is isolated from the rest by the dry Chicamocha Canyon while cluster J is separated from the other NA clusters by the Marañón Valley (Fig. 1). Särkinen et al. (2012) suggested that biome heterogeneity across the Andes represented a strong barrier to dispersal within island-like ecosystems. This is particularly relevant when deep valleys segment the mountain ranges, as is the case here. In addition, for *O. venezuelensis*, clusters H and I have ITS haplotypes distinct from others in the species, namely Hn28 and Hn30 (Fig. 1, Supp. Fig. 1). These haplotypes are distributed in the southernmost part of these species' distribution range and their differentiation from species-specific haplotypes distributed in the northernmost areas (Hn26, Hn27 and Hn28) further supports the observed phylogeographic structure and possible pattern of isolation by distance.

Genetic patterns in the light of Quaternary glacial-interglacial cycles.

Our dated tree (Fig. 2) is consistent with Quaternary diversification in the NAC, and high levels of molecular diversity for both nuclear and plastid regions, as well as the high number of unsampled cpDNA haplotypes in our dataset, are concordant with a scenario of expansion and contraction of Páramo islands during the glacial cycles of the Quaternary (Table 4, Supp. Table 8 and Supp. Fig. 2). SAMOVA analysis failed to identify any clear groupings within the NAC (Supp. Table 8, Supp. Figs. 1 – 2) and variation amongst NA clusters was moderate and mostly explained by within cluster variation (ITS, 86%; cpDNA, 79%; Table 3). Vicariance events would allow

for differentiation of populations and diversification, through selection and drift. If reproductive isolation is incomplete, subsequent expansion events may have allowed gene flow amongst nearby populations and potentially even amongst species. Repeated vicariance and contact, which would be expected from Quaternary glacial cycles, would generate complex genetic patterns, with species sharing haplotypes. Such patterns are evident in *Oreobolus*, with a few widespread haplotypes amongst species apparently giving rise to geographically restricted haplotypes (Supp. Figs. 1 – 2). Similar patterns have been reported for the afro-alpine populations of *Arabis alpina* where several cycles of range contraction and expansion caused by the glacial cycles of the Quaternary may have shaped intra-specific distribution of genetic diversity (Assefa et al. 2007). In the same way, cluster M in the SA region is a divergent genetic group for both ITS and cpDNA in SAMOVA analyses (Supp. Figs. 1 – 2). Molecular diversity indices for this cluster showed low haplotype diversity and high nucleotide diversity in ITS, and high haplotype diversity and low nucleotide diversity in cpDNA (Supp. Table 8). A possible explanation for this pattern might be that these populations underwent a bottleneck during isolation resulting in a low number of divergent haplotypes. During the glacial cycles of the Quaternary ice sheets covered extensive areas and generated massive fragmentation and restriction in the distribution of southern Andean plants producing pockets of refugial populations (e.g. Markgraf et al. 1995). Although a scenario of Pleistocene refugia has already been proposed for other southern Andean plants (e.g. Tremetsberger et al. 2009) further work would be required to assess the potential for refugial populations in *O. obtusangulus* (SA).

Glacial cycles may have also had an impact at the inter-specific level. *Oreobolus ecuadorensis* has the lowest molecular diversity indices for both ITS and cpDNA (Table 1) and is one of the most geographically restricted species, found only in Ecuador and northern Peru (Fig. 1). Such patterns may arise through a severe bottleneck followed by a population expansion likely imposed by the glacial cycles of the Quaternary (Templeton 1998; Hewitt 2004). Ecuador and Peru have the highest percentage of permanent snow and therefore interglacial periods may have greatly reduced the size of the populations of *O. ecuadorensis*, reducing its genetic diversity. Following the Last Glacial Maximum (LGM), population expansion may have occurred with new mutations likely to accumulate as the species occupied new areas. New haplotypes were thereby produced, diverging from the founder population by only a few nucleotides. At the same time, the strong impact of interglacial periods is evident in the clear differentiation of *O. ecuadorensis* from all other species (Table 2, Supp. Figs. 3 – 4).

There was no clear evidence of ongoing hybridization but historic hybridization between sympatric sister species *O. cleefii* and *O. obtusangulus* (NA) may have been facilitated by periods of isolation and divergence during the glacial cycles of the Quaternary. Secondary contact zones can form from long-distance dispersal events, leading to interspecific hybridization, such as that proposed by Gizaw et al. (2016) for two co-occurring sister species of *Carex* from a similar tropical alpine ecosystem in East Africa. We suggest a similar scenario for *O. cleefii* and *O. obtusangulus* (NA), with renewed contact occurring following isolation during interglacial periods in the Quaternary (van der Hammen 1974).

CONCLUSION

This is one of a few studies to investigate genetic relationships both within and between species in a recently diverged Páramo genus and hence it provides a significant contribution to the understanding of the historical assembly of the Páramo flora. The results presented here are consistent with a role for contraction and expansion of Páramo islands during glacial cycles in the diversification of *Oreobolus* species. ILS appears to have played a role in the complex genetic patterns observed amongst these recently diverged *Oreobolus* species. ILS rather than recent hybridization is suggested by the lack of heterozygosity in ITS, but a role for historical hybridization cannot be discounted, particularly in several situations where the species are sympatric. Additional work incorporating more extensive sampling of individuals and assessing additional genetic data will be required to more accurately estimate patterns of historical demography of *Oreobolus*, which could bring further insight into the population dynamics of Páramo plants.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

DECLARATION OF AUTHORSHIP

MCGG and JER devised the project. LEN assisted with data analyses. MCGG drafted the text, with substantial contributions by JER, RTP and LEN. All authors contributed to final editing.

REFERENCES

Abbott R, Albach D, Ansell S, et al (2013) Hybridization and speciation. J Evol Biol 26:229–246. doi: 10.1111/j.1420-9101.2012.02599.x

Arnold ML (1997) Natural hybridization and evolution. Oxford University Press, New York

Assefa A, Ehrich D, Taberlet P, et al (2007) Pleistocene colonization of afro-alpine "sky islands" by the arctic-alpine *Arabis alpina*. Heredity 99:133–142. doi: 10.1038/sj.hdy.6800974

Bouckaert R, Drummond AJ (2017) bModelTest: Bayesian phylogenetic site model averaging and model comparison. BMC Evolutionary Biology 17:1–11. doi: 10.1186/s12862-017-0890-6

Bouckaert R, Heled J, Kühnert D, et al (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput Biol 10:e1003537–6. doi: 10.1371/journal.pcbi.1003537

Britton MN, Hedderson TA, Verboom GA (2014) Topography as a driver of cryptic speciation in the high-elevation cape sedge *Tetraria triangularis* (Boeck.) C. B. Clarke (Cyperaceae: Schoeneae). Molecular Phylogenetics and Evolution 77:96–109. doi: 10.1016/j.ympev.2014.03.024

Bryant D, Moulton V (2004) Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. Molecular Biology and Evolution 21:255–265. doi: 10.1093/molbev/msh018

Burban C, Petit R, Carcreff E, Jactel H (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. Mol Ecol 8:1593–1602.

Buytaert W, Cuesta-Camacho F, Tobón C (2010) Potential impacts of climate change on the environmental services of humid tropical alpine regions. Global Ecology and Biogeography 20:19–33. doi: 10.1111/j.1466-8238.2010.00585.x

Chacón J, Madriñán S, Chase MW, Bruhl JJ (2006) Molecular phylogenetics of *Oreobolus* (Cyperaceae) and the origin and diversification of the American species. Taxon 55:359–366.

Cutter AD (2013) Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. Molecular Phylogenetics and Evolution 69:1172–1185. doi: 10.1016/j.ympev.2013.06.006

Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution 24:332–340. doi: 10.1016/j.tree.2009.01.009

Doyle J, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15.

Drummond AJ, Bouckaert RR (2015) Bayesian Evolutionary Analysis with BEAST, First. Cambridge University Press, Cambridge

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571–2581.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10:564–567. doi: 10.1111/j.1755-0998.2010.02847.x

French GC (2003) Conservation genetics of British *Euphrasia* L. University of Edinburgh and Royal Botanic Garden Edinburgh, Edinburgh

Gizaw A, Kebede M, Nemomissa S, et al (2013) Phylogeography of the heathers *Erica arborea* and *E. trimera* in the afro-alpine “sky islands” inferred from AFLPs and plastid DNA sequences. Flora 208:453–463. doi: 10.1016/j.flora.2013.07.007

Gizaw A, Wondimu T, Mugizi TF, et al (2016) Vicariance, dispersal, and hybridization in a naturally fragmented system: the afro-alpine endemics *Carex monostachya* and *C. runssoroensis* (Cyperaceae). Alpine Botany 126:59–71. doi: 10.1007/s00035-015-0162-2

Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F-

statistics. Molecular Ecology Notes 5:184–186. doi: doi: 10.1111/j.1471-8278.2004.00828.x

Gómez-Gutiérrez MC (2016) Evolution in the high-altitude Páramo ecosystem. The University of Edinburgh PhD thesis.

Graham A (2009) The Andes: a geological overview from a biological perspective. Annals of the Missouri Botanical Garden 96:371–385. doi: 10.3417/2007146

Griffin PC, Hoffmann AA (2014) Limited genetic divergence among Australian alpine *Poa* tussock grasses coupled with regional structuring points to ongoing gene flow and taxonomic challenges. Annals of Botany 113:953–965. doi: 10.1093/aob/mcu017

Heled J, Drummond AJ (2012) Calibrated Tree Priors for Relaxed Phylogenetics and Divergence Time Estimation. Systematic Biology 61:138–149. doi: 10.1093/sysbio/syr087

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society B: Biological Sciences 359:183–195. doi: 10.1098/rstb.2003.1388

Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Systematic Biology 58:367–380. doi: 10.1093/sysbio/syp035

Hooghiemstra H, van der Hammen T (2004) Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. Philosophical Transactions of the Royal Society B: Biological Sciences 359:173–181. doi: 10.1098/rstb.2003.1420

Hooghiemstra H, Wijninga VM, Cleef AM (2006) The paleobotanical record of Colombia: implications for biogeography and biodiversity. Annals of the Missouri Botanical Garden 93:297–324. doi: 10.3417/0026-6493(2006)93[297:TPROCI]2.0.CO;2

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23:254–267. doi: 10.1093/molbev/msj030

Jakob SS, Blattner FR (2006) A chloroplast genealogy of *Hordeum* (Poaceae): long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. Molecular Biology and Evolution 23:1602–1612. doi: 10.1093/molbev/msl018

Kadu CAC, Konrad H, Schueler S, et al (2013) Divergent pattern of nuclear genetic diversity across the range of the Afromontane *Prunus africana* mirrors variable climate of African highlands. Annals of Botany 111:47–60. doi: 10.1093/aob/mcs235

Kebede M, Ehrich D, Taberlet P, et al (2007) Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. Mol Ecol 16:1233–1243. doi: 10.1111/j.1365-294X.2007.03232.x

Kolář F, Dušková E, Sklenář P (2016) Niche shifts and range expansions along cordilleras drove diversification in a high-elevation endemic plant genus in the tropical Andes. Mol Ecol 25:4593–4610. doi: 10.1111/mec.13788

Luteyn JL (1999) Páramos: a checklist of plant diversity, geographical distribution, and botanical literature, First. The New York Botanical Garden Press, New York

Maddison WP, Knowles LL (2006) Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55:21–30. doi: 10.1080/10635150500354928

Maddison WP, Maddison DR (2014) Mesquite: a modular system for evolutionary analysis. Version 2.75. Available at http://mesquiteproject.org. – mesquiteproject.org.

Madriñán S, Cortés AJ, Richardson JE (2013) Páramo is the world's fastest evolving and coolest biodiversity hotspot. Frontiers in Genetics 4:1–7. doi: 10.3389/fgene.2013.00192/abstract

Markgraf V, McGlone M, Hope G (1995) Neogene paleoenvironmental and paleoclimatic change in southern temperate ecosystems—a southern perspective. Trends in Ecology & Evolution 10:143–147. doi: 10.1016/S0169-5347(00)89023-0

Moore DM (1967) Chromosome numbers of Falkland Islands Angiosperms. British Antarctic Survey Bulletin 14:69–82.

Mora-Osejo LE (1987) Estudios morfológicos, autoecológicos y sistemáticos en Angiospermas. Academia Colombiana de Ciencias Exactas, Físicas y Naturales, Bogotá, DE

Möller M, Cronk QCB (1997) Origin and relationships of *Saintpaulia* (Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. American Journal of Botany 84:956–965.

Naciri Y, Linder HP (2015) Species delimitation and relationships: The dance of the seven veils. Taxon 64:3–16. doi: 10.12705/641.24

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144:1237–1245.

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/. httpswww.R-project.org

Rambaut A, Drummond AJ (2015) TreeAnnotator: MCMC output analysis. Version

2.3.0. Available at http://tree.bio.ed.ac.uk/software/beast. httptree.bio.ed.ac.uksoftware

Rambaut A, Suchard MA, Xie W, Drummond AJ (2013) Tracer: MCMC trace analysis tool. Version 1.6.0. Available at http://tree.bio.ed.ac.uk/software/tracer. httptree.bio.ed.ac.uksoftware

Sang T, Crawford D, Stuessy T (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). American Journal of Botany 84:1120–1136.

Särkinen T, Pennington RT, Lavin M, et al (2012) Evolutionary islands in the Andes: persistence and isolation explain high endemism in Andean dry tropical forests. Journal of Biogeography 39:884–900. doi: 10.1111/j.1365-2699.2011.02644.x

Schaal BA, Hayworth DA, Olsen KM, et al (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7:465–474. doi: 10.1046/j.1365-294x.1998.00318.x

Schaal BA, Olsen KM (2000) Gene genealogies and population variation in plants. Proc Natl Acad Sci USA 97:7024–7029.

Seberg O (1988) Taxonomy, phylogeny, and biogeography of the genus *Oreobolus* R.Br. (Cyperaceae), with comments on the biogeography of the South Pacific continents. Botanical Journal of the Linnean Society 96:119–195. doi: 10.1111/j.1095-8339.1988.tb00632.x

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in Angiosperms: the tortoise and the hare III. American Journal of Botany 94:275–288. doi: 10.3732/ajb.94.3.275

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. Systematic Biology 49:369–381.

Simpson BB (1975) Pleistocene changes in the flora of the high tropical Andes. Paleobiology 1:273–294. doi: 10.2307/2400369

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109.

Tate JA, Simpson BB (2003) Paraphyly of Tarasa (Malvaceae) and diverse origins of the polyploid species. Systematic Botany 28:723–737. doi: 10.2307/25063919?ref=search-gateway:ab20df06ffd5b47899780d1c7b361d21

Teacher AGF, Griffiths DJ (2011) HapStar: automated haplotype network layout and visualization. Molecular Ecology Resources 11:151–153. doi: 10.1111/j.1755-0998.2010.02890.x

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing

hypotheses about gene flow and population history. Mol Ecol 7:381–397.

Tremetsberger K, Urtubey E, Terrab A, et al (2009) Pleistocene refugia and polytopic replacement of diploids by tetraploids in the Patagonian and Subantarctic plant *Hypochaeris incana* (Asteraceae, Cichorieae). Mol Ecol 18:3668–3682. doi: 10.1111/j.1365-294X.2009.04298.x

van der Hammen T (1974) The Pleistocene changes of vegetation and climate in tropical South America. Journal of Biogeography 1:3–26.

van der Hammen T, Cleef AM (1986) Development of the high Andean Páramo flora and vegetation. In: Vuilleumier F, Monasterio M (eds) High Altitude Tropical Biogeography, First. Oxford University Press, Inc, New York, pp 153–201

Vásquez DLA, Balslev H, Hansen MM, et al (2016) Low genetic variation and high differentiation across sky island populations of *Lupinus alopecuroides* (Fabaceae) in the northern Andes. Alpine Botany 1–8. doi: 10.1007/s00035-016-0165-7

Wondimu T, Gizaw A, Tusiime FM, et al (2013) Crossing barriers in an extremely fragmented system: two case studies in the afro-alpine sky island flora. Plant Syst Evol 300:415–430. doi: 10.1007/s00606-013-0892-9

TABLES

Table 1. Molecular diversity indices for ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L) for each species. N: number of individuals; H: number of haplotypes; hr: haplotype richness (ITS, rarefied to a minimum sample of 15; cpDNA, rarefied to a minimum sample of 9); h, haplotype diversity (± SD); π, nucleotide diversity (± SD). A, *O. obtusangulus* considered as one species; B, *O. obtusangulus* considered as two species.

Species	N	H	hr	h	π x 100
ITS					
O. cleefii	15	5	5.00	0.70 ± 0.11	0.45 ± 0.30
O. ecuadorensis	24	4	3.12	0.31 ± 0.12	0.01 ± 0.10
O. goeppingeri	75	12	6.09	0.79 ± 0.03	1.15 ± 0.61
O. obtusangulus					
NA	23	8	6.12	0.68 ± 0.10	0.56 ± 0.35
SA	33	5	3.88	0.64 ± 0.06	2.25 ± 1.16
Combined	56	13	6.59	0.82 ± 0.03	2.76 ± 1.39
O. venezuelensis	27	7	5.02	0.63 ± 0.10	1.49 ± 0.80
cpDNA					
O. cleefii	9	4	4.00	0.78 ± 0.11	1.96 ± 1.07
O. ecuadorensis	29	5	3.54	0.72 ± 0.05	0.11 ± 0.07
O. goeppingeri	27	11	5.67	0.84 ± 0.06	2.36 ± 1.17
O. obtusangulus					
NA	20	10	6.35	0.91 ± 0.04	1.70 ± 0.86
SA	19	8	5.22	0.84 ± 0.06	2.20 ± 1.11
Combined	39	18	7.12	0.94 ± 0.02	3.05 ± 1.49
O. venezuelensis	14	8	6.30	0.91 ± 0.05	2.23 ± 1.15

Table 2. Pairwise F_{ST} values amongst species calculated from ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L) considering *O. obtusangulus* as (a) one species and (b) as two species. Values for ITS are below the diagonal and cpDNA above. Bold numbers denote significance at the 5% level. cle: *O. cleefii*, ecu: *O. ecuadorensis*, goe: *O. goeppingeri*, obt: *O. obtusangulus* and ven: *O. venezuelensis*.

(a)

	cle	ecu	goe	obt	ven	
cle		**0.797**	**0.283**	**0.098**	**0.317**	cle
ecu	**0.770**		**0.732**	**0.600**	**0.801**	ecu
goe	**0.284**	**0.307**		**0.229**	**0.288**	goe
obt	**0.269**	**0.360**	**0.289**		**0.256**	obt
ven	**0.314**	**0.328**	**0.175**	**0.291**		ven
	cle	ecu	goe	obt	ven	

(b)

	cle	ecu	goe	obt (NA)	obt (SA)	ven	
cle		**0.797**	**0.283**	-0.020	**0.487**	**0.317**	cle
ecu	**0.770**		**0.732**	**0.780**	**0.819**	**0.801**	ecu
goe	**0.284**	**0.307**		**0.363**	**0.430**	**0.288**	goe
obt (NA)	**0.157**	**0.710**	**0.294**		**0.547**	**0.399**	obt (NA)
obt (SA)	**0.595**	**0.649**	**0.578**	**0.620**		**0.478**	obt (SA)
ven	**0.314**	**0.328**	**0.175**	**0.339**	**0.551**		ven
	cle	ecu	goe	obt (NA)	obt (SA)	ven	

Table 3. Analysis of molecular variance (AMOVA) results for ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L).

Group level	Source of variation	Degrees of freedom		Sum of Squares		Variance components		Percentage of variation		Fixation indices	
		ITS	cpDNA	ITS	cpDNA	ITS	cpDNA	ITS	cpDNA	ITS	cpDNA
Species	Among species	4	4	260	2033	1.69	21.67	30.46	47.65	F_{ST} = 0.31***	F_{ST} = 0.48***
	Within species	192	113	741	2689	3.86	23.80	69.54	52.35		
Clusters (all clusters)	Among clusters	13	12	442	1925	2.30	15.54	42.95	36.84	F_{ST} = 0.43***	F_{ST} = 0.37***
	Within clusters	183	105	560	2798	3.06	26.65	57.05	63.16		
Clusters (northern Andes - NA)	Among clusters	9	8	88	885	0.46	7.88	14.50	21.34	F_{ST} = 0.15**	F_{ST} = 0.21***
	Within clusters	154	90	416	2612	2.70	29.03	85.50	78.66		
Continental regions (NA vs SA)	Among regions	1	1	309	746	5.38	19.89	59.49	35.54	F_{CT} = 0.60***	F_{CT} = 0.36***
	Among clusters within regions	12	11	133	1179	0.61	9.41	6.72	16.82	F_{SC} = 0.17**	F_{SC} = 0.26***
	Within clusters	183	105	560	2798	3.06	26.65	33.79	47.64	F_{ST} = 0.66***	F_{ST} = 0.52***
SAMOVA groups	Among groups	2	2	348	999	5.71	25.47	62.19	42.59	F_{CT} = 0.62***	F_{CT} = 0.43***
	Among clusters within groups	11	10	93	926	0.41	7.68	4.51	12.84	F_{SC} = 0.12*	F_{SC} = 0.22***
	Within clusters	183	105	560	2798	3.06	26.65	33.30	44.56	F_{ST} = 0.67***	F_{ST} = 0.55***

* significant at the 5% level; ** significant at the 1% level; *** significant at the 0.1% level

Table 4. Molecular diversity indices for ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L) for each SAMOVA grouping. N: number of individuals; H: number of haplotypes; hr: haplotype richness (ITS, rarefied to a minimum sample of 16; cpDNA, rarefied to a minimum sample of 9); h, haplotype diversity (± SD); π, nucleotide diversity (± SD).

SAMOVA group	N	H	hr	h	π x 100
	ITS				
I	164	25	9.20	0.91 ± 0.01	1.18 ± 0.63
II	17	5	4.88	0.58 ± 0.13	2.18 ± 1.17
III	16	2	2.00	0.13 ± 0.11	1.46 ± 0.80
	cpDNA				
I	100	33	7.43	0.95 ± 0.01	3.03 ± 1.46
II	9	4	4.00	0.58 ± 0.18	1.90 ± 1.04
III	9	4	4.00	0.75 ± 0.11	0.06 ± 0.05

FIGURES

Fig. 1 Geographical distribution of *Oreobolus* in South America based on herbarium records (coloured dots). Sampling localities (1 – 32) and their corresponding cluster (A – N) are also indicated. Arrows denote geographical features.

Fig. 2 Maximum clade credibility tree from the *BEAST 2 analysis based on ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L). Numbers above the branches represent posterior probability values. Node bars show 95% HPD. NAC, northern Andean clade.

Fig. 3 NeighborNet network for the ITS haplotypes based on the uncorrected-p distances. Haplotypes are coloured according to species. Shared haplotypes are shown in white, with pie charts below (labelled with haplotype number) showing the frequency per species. NA: northern Andes, SA: southern Andes

Fig. 4 NeighborNet network for the cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L) haplotypes based on the uncorrected-p distances. Haplotypes are coloured according to species. Shared haplotypes are shown in white, with pie charts (labelled with haplotype number) indicating frequency per species shown below. NA: northern Andes, SA: southern Andes

Fig. 5 NeighborNet network showing genetic relatedness amongst clusters based on ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L) F_{ST} pairwise values.

ELECTRONIC SUPPLEMENTARY MATERIAL

Supplementary Table 1. Geographic coordinates and corresponding cluster of the sampling localities.

Nº	SAMPLING LOCALITY	CLUSTER	LATITUDE	LONGITUDE
1	CHIRRIPO	A	9.48411000	-83.48861000
2	COCUY	B	6.41211667	-72.33128333
3	LA RUSIA	C	5.93951667	-73.07583333
4	IGUAQUE	C	5.68610000	-73.44773333
5	TOTA-BIJAGUAL	B	5.48143333	-72.85540000
6	RABANAL	C	5.40818333	-73.54915000
7	GUERRERO	C	5.22618333	-74.01788333
8	CHINGAZA	D	4.52848333	-73.75866667
9	SUMAPAZ	D	4.28958333	-74.20781667
10	PURACE	E	2.36088333	-76.35038333
11	AZUFRAL	F	1.09543333	-77.68711667
12	VOLCAN CHILES	F	0.80000000	-77.93333333
13	MIRADOR	F	0.56666667	-77.65000000
14	COTOCACHI	F	0.36666667	-78.33333333
15	COTOPAXI	G	-0.66666667	-78.36666667
16	LLANGANATI	G	-1.15000000	-78.30000000
17	ALAO-HUAMBOYA	G	-1.80000000	-78.43333333
18	PARAMO DE LAS CAJAS	H	-2.81666667	-79.26666667
19	CUENCA-LIMON	H	-3.00000000	-78.66666667
20	CUENCA-LOJA	H	-3.16666667	-79.03333333
21	PODOCARPUS	I	-4.40000000	-79.10000000
22	CAJAMARCA	J	-7.05000000	-78.58333333
23	HUASCARAN	J	-9.45000000	-77.26666000
24	VALDIVIA	K	-40.18333333	-73.51666666
25	FIORDO PEEL	L	-50.50000000	-73.73333333
26	MALVINAS	N	-51.64297000	-59.89473000
27	MORRO PHILIPPI	L	-51.73333333	-71.50000000
28	MAGALLANES	L	-53.45000000	-71.76666700
29	TIERRA DEL FUEGO	M	-54.76666666	-67.40000000
30	ISLA DE LOS ESTADOS	M	-54.80000000	-64.31666666
31	ISLA NAVARINO	M	-55.07553100	-67.65539600
32	CABO DE HORNOS	M	-55.94407800	-67.28092500

Supplementary Table 2. Sequence information

Supplementary Table 3. Number of individuals successfully sequenced per species per sampling locality for ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L). Areas where species are not distributed are noted as n.d.

CLUSTER/Sampling locality	*O. cleefii*		*O. ecuadorensis*		*O. goeppingeri*		*O. obtusangulus*		*O. venezuelensis*	
	ITS	cpDNA	ITS	cpDNA	ITS	cpDNA	ITS	cpDNA	ITS	cpDNA
CLUSTER A										
(1) Chirripo	n.d.	n.d.	n.d.	n.d.	2	-	n.d.	n.d.	-	-
CLUSTER B										
(2) Cocuy	5	4	n.d.	n.d.	3	4	-	-	-	-
(5) Tota-Bijagual	2	1	n.d.	n.d.	2	1	-	-	-	-
CLUSTER C										
(4) Iguaque	-	-	n.d.	n.d.	1	1	-	-	-	-
(3) La Rusia	2	2	n.d.	n.d.	1	-	-	-	2	1
(6) Rabanal	-	-	n.d.	n.d.	2	1	-	-	-	-
(7) Guerrero	1	-	n.d.	n.d.	1	-	-	-	-	-
CLUSTER D										
(8) Chingaza	1	-	n.d.	n.d.	3	1	-	-	2	1
(9) Sumapaz	-	-	n.d.	n.d.	3	2	1	-	4	2
CLUSTER E										
(10) Purace	n.d.	n.d.	n.d.	n.d.	3	3	-	-	-	-
CLUSTER F										
(11) Azufral	4	2	-	-	1	1	-	-	-	-
(12) Volcan Chiles	-	-	1	1	5	-	5	4	-	-
(13) Mirador	-	-	-	-	2	2	1	2	1	1
(14) Cotocachi	n.d.	n.d.	1	2	3	2	-	-	-	-
CLUSTER G										
(15) Cotopaxi	n.d.	n.d.	9	13	2	-	2	2	-	1
(16) Llanganati	n.d.	n.d.	-	1	2	-	2	1	-	-

(17) Alao-Huamboya	n.d.	n.d.	3	2	3	-	-	-	-	-
CLUSTER H										
(18) Paramo De Las Cajas	n.d.	n.d.	3	4	2	2	4	3	-	-
(19) Cuenca-Limon	n.d.	n.d.	-	-	2	-	3	3	-	-
(20) Cuenca-Loja	n.d.	n.d.	4	4	11	3	3	3	2	3
CLUSTER I										
(21) Podocarpus	n.d.	n.d.	-	-	18	4	1	1	15	4
CLUSTER J										
(22) Cajamarca	n.d.	n.d.	1	1	3	-	1	1	-	-
(23) Huascaran	n.d.	n.d.	2	1	-	-	-	-	-	-
CLUSTER K										
(24) Valdivia	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	1	1	n.d.	n.d.
CLUSTER L										
(25) Fiordo Peel	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	2	-	n.d.	n.d.
(27) Morro Philippi	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	2	1	n.d.	n.d.
(28) Magallanes	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	11	7	n.d.	n.d.
CLUSTER M										
(29) Tierra Del Fuego	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	10	5	n.d.	n.d.
(30) Isla De Los Estados	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	2	1	n.d.	n.d.
(31) Isla Navarino	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	1	-	n.d.	n.d.
(32) Cabo De Hornos	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	3	3	n.d.	n.d.
CLUSTER N										
(26) Malvinas	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	1	1	n.d.	n.d.
TOTAL	15	9	24	29	75	27	56	39	27	14

Supplementary Table 4. Frequency of occurrence of ITS haplotypes (Hn) across clusters and species. Clusters (A – N) as described in Figure 1 and Supplementary Table 1. cle: *O. cleefii*, ecu: *O. ecuadorensis*, goe: *O. goeppingeri*, obt: *O. obtusangulus* and ven: *O. venezuelensis*.

		NORTHERN ANDES										SOUTHERN ANDES			
		A	B	C	D	E	F	G	H	I	J	K	L	M	N
Hn1	*cle*	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn2	*cle*	.	7	1	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn3	*cle*	.	.	2	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	2	11	8	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	1	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn4	*cle*	.	.	.	.	.	3	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	2	3	.	1	.	.	.	.	.
	obt	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn5	*cle*	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	2	1	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn6	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn7	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	2	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn8	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	1	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	1	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn9	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	2	6	3	5	13	.	.	.	.	.
	obt	.	.	.	.	.	.	.	1	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn10	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	1	1	3	1	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	ven	.	1	1	2	.	1	.	.	.	.	.	.	.	.
Hn11	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	2	4	3	.	3	.	3	.	1	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn12	*goe*	.	.	.	.	.	.	.	7	2	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn13	*goe*	.	.	.	.	.	.	.	.	1	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn14	*goe*	.	.	.	.	.	.	.	.	1	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn15	*goe*	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn16	*goe*	.	2	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	2	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn17	*goe*	.	.	.	.	.	.	.	.	.	1	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn18	*goe*	.	.	.	.	.	.	.	.	.	1	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn19	*goe*	2	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn20	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	10	.	1
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn21	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	1	1	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn22	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	2	15	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn23	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	1	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn24	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	3	2	8	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn25	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	2	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn26	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	1	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn27	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn28	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	1	15	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn29	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hn30	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	1	.	.	.	.	.	.

Supplementary Table 5. Frequency of occurrence of cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L) haplotypes (Hc) across clusters and species. Clusters (B – N) as described in Figure 1 and Supplementary Table 1. cle: *O. cleefii*, ecu: *O. ecuadorensis*, goe: *O. goeppingeri*, obt: *O. obtusangulus* and ven: *O. venezuelensis*.

		NORTHERN ANDES									SOUTHERN ANDES			
		B	C	D	E	F	G	H	I	J	K	L	M	N
Hc1	*cle*	.	2	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc2	*cle*	4	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	3	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc3	*cle*	1	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc4	*cle*	.	.	.	.	2	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	1	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc5	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	3	9	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc6	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	2	5	.	1	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	1	1	1	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc7	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	4	3	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc8	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	1	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc9	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	1	.	.	.	.
	goe	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc10	*cle*	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
	goe	.	.	.	.	.	.	.	1	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.

	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc11	*goe*	.	.	.	.	2	.	2	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc12	*goe*	2	1	2	.	2	.	3	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc13	*goe*	1	.	1	1	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	1	2	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc14	*goe*	2	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	1	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc15	*goe*	.	.	.	1	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc16	*goe*	.	.	.	1	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc17	*goe*	.	.	.	.	1	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc18	*goe*	.	1	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc19	*goe*	.	.	.	.	.	.	.	2	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc20	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	5	.	1
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc21	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	1	4	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc22	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	1	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.

	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc23	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	1	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc24	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	ven	.	.	.	.	1	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc25	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc26	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	1	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc27	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	1	2	2	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc28	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc29	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc30	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc31	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	2	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc32	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	2	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc33	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	3	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc34	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	1	.	.	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc35	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	1	.
	ven	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc36	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc37	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	1	1	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc38	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	2	.	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc39	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	3	.	.	.	.	.	.
	cle	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ecu	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hc40	*goe*	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	obt	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ven	.	.	.	.	.	.	.	1	.	.	.	.	.	.

Supplementary Table 6. Spatial analysis of molecular variance (SAMOVA) results for ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32*-*trn*L) showing the variance amongst groups (F_{CT} values) for pre-defined K number of groups.

	K											
	2	3	4	5	6	7	8	9	10	11	12	13
F_{CT} ITS	0.595	0.622	0.608	0.608	0.603	0.581	0.505	0.507	0.468	0.481	0.504	0.639
F_{CT} cpDNA	0.417	0.426	0.417	0.414	0.412	0.406	0.405	0.410	0.441	0.502	0.675	-

Supplementary Table 7. Pairwise F_{ST} values amongst clusters calculated from ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32-trn*L). Results for ITS are shown below the diagonal and cpDNA above. Bold numbers indicate significance at the 5% level.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
A		-	-	-	-	-	-	-	-	-	-	-	-	-	A
B	-		0.134	0.112	0.136	0.001	**0.504**	**0.101**	**0.211**	**0.586**	-	**0.462**	**0.713**	-	B
C	-	**0.092**		0.028	0.050	0.038	**0.206**	-0.003	**0.157**	0.200	-	**0.417**	**0.690**	-	C
D	-	**0.141**	0.050		0.023	**0.166**	**0.530**	**0.220**	**0.264**	**0.597**	-	**0.470**	**0.740**	-	D
E	-	**0.290**	-0.073	0.200		0.048	**0.511**	0.113	0.076	0.712	-	**0.414**	**0.874**	-	E
F	-	**0.147**	**0.106**	**0.201**	0.101		**0.342**	0.000	**0.109**	**0.402**	-	**0.408**	**0.621**	-	F
G	-	**0.267**	0.061	**0.249**	0.206	**0.089**		**0.188**	**0.452**	-0.080	-	**0.646**	**0.794**	-	G
H	-	**0.258**	**0.120**	**0.258**	0.028	0.051	0.052		**0.095**	**0.207**	-	**0.406**	**0.581**	-	H
I	-	**0.232**	0.055	**0.202**	-0.046	**0.141**	**0.127**	**0.087**		**0.532**	-	**0.444**	**0.724**	-	I
J	-	**0.498**	**0.065**	**0.484**	**0.485**	**0.406**	**0.232**	**0.276**	**0.122**		-	**0.698**	**0.989**	-	J
K	-	-	-	-	-	-	-	-	-	-		-	-	-	K
L	-	**0.629**	**0.474**	**0.635**	**0.482**	**0.659**	**0.660**	**0.661**	**0.537**	**0.595**	-		**0.657**	-	L
M	-	**0.702**	**0.531**	**0.708**	**0.601**	**0.715**	**0.718**	**0.704**	**0.563**	**0.688**	-	**0.296**		-	M
N	-	-	-	-	-	-	-	-	-	-	-	-	-		N
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	

Supplementary Table 8. Molecular diversity indices for ITS and cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl32*-*trn*L) for each cluster. Clusters (A – N) as described in Figure 1 and Supplementary Table 1. Metrics were not applicable (n.a.) for clusters with less than three individuals. N, number of individuals; h, haplotype diversity (± SD); π, nucleotide diversity (± SD).

	ITS				cpDNA			
	SAMOVA group	N	h	π x100	SAMOVA group	N	h	π x 100
A	I	2	n.a.	n.a.	-	-	-	-
B	I	13	0.69 ± 0.12	0.76 ± 0.46	I	10	0.82 ± 0.10	2.34 ± 1.25
C	I	10	0.82 ± 0.10	2.95 ± 1.63	I	5	0.90 ± 0.16	3.95 ± 2.41
D	I	14	0.85 ± 0.07	0.63 ± 0.39	I	6	0.73 ± 0.16	2.78 ± 1.62
E	I	3	0.67 ± 0.31	0.72 ± 0.63	I	3	1.00 ± 0.27	2.15 ± 1.62
F	I	24	0.86 ± 0.04	0.67 ± 0.40	I	17	0.93 ± 0.04	2.68 ± 1.36
G	I	23	0.76 ± 0.08	0.66 ± 0.40	I	20	0.77 ± 0.08	1.69 ± 0.86
H	I	34	0.83 ± 0.03	0.84 ± 0.48	I	25	0.92 ± 0.03	3.10 ± 1.55
I	I	34	0.67 ± 0.05	1.88 ± 0.98	I	10	0.91 ± 0.08	2.20 ± 1.18
J	I	7	0.91 ± 0.10	0.56 ± 0.39	I	3	0.67 ± 0.31	0.09 ± 0.08
K	II	1	n.a.	n.a.	I	1	n.a.	n.a.
L	II	15	0.55 ± 0.14	2.36 ± 1.27	II	8	0.64 ± 0.18	2.13 ± 1.18
M	III	16	0.13 ± 0.11	1.46 ± 0.80	III	9	0.75 ± 0.11	0.06 ± 0.05
N	II	1	n.a.	n.a.	II	1	n.a.	n.a.

Supplementary Fig. 1 MST and distribution of ITS haplotypes. Numbers refer to haplotypes listed in Supplementary Table 5. Haplotypes are coloured according to species. Shared haplotypes are shown in white. Detail of species sharing haplotypes is given in Fig. 3. Hypothetical haplotypes are represented by filled black circles. Letters on the map refer to clusters as described in Figure 1 and Supplementary Table 3. Pie charts are proportional to sample size for each cluster (N = 1 – 34). Numbers next to each segment refer to haplotype number. NA: northern Andes, SA: southern Andes

Supplementary Fig. 2 MST and distribution of cpDNA (*trn*L-F, *trn*H-*psb*A and *rpl*32-*trn*L) haplotypes. Numbers refer to haplotypes listed in Supplementary Table 6. Haplotypes are coloured according to species. Shared haplotypes are shown in white. Detail of species sharing haplotypes is given in Fig. 5. Hypothetical haplotypes are represented by filled black circles, numbers within indicate their number when more than one. Letters on the map refer to clusters as described in Figure 1 and Supplementary Table 3. Pie charts are proportional to sample size for each cluster (N = 1 – 25). Numbers next to each segment refer to haplotype number. NA: northern Andes, SA: southern Andes

Supplementary Fig. 3 NeighborNet network showing genetic relatedness amongst the South American species of *Oreobolus* based on ITS F_{ST} pairwise values considering (a) *O. obtusangulus* as one species (b) *O. obtusangulus* as two species

Supplementary Fig. 4 NeighborNet network showing genetic relatedness amongst the South American species of *Oreobolus* based on cpDNA (*trn*L-F, *trn*H-*psb*A and

*rpl32-trn*L) F_{ST} pairwise values considering (a) *O. obtusangulus* as one species (b) *O. obtusangulus* as two species

Figure 1

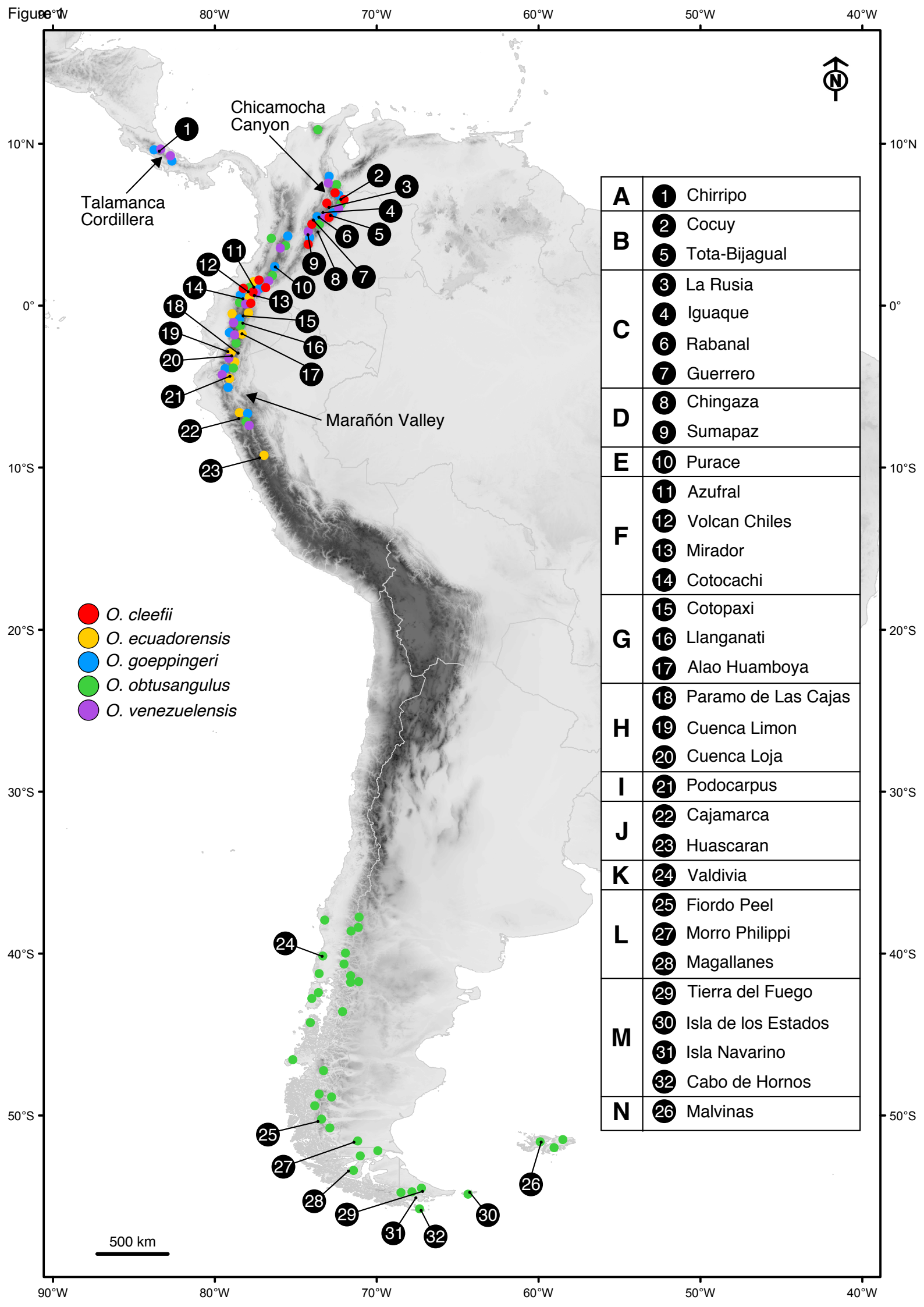

90°W
80°W
70°W
60°W
50°W
40°W
10°N
0°
10°S
20°S
30°S
40°S
50°S
Chicamocha Canyon
Talamanca Cordillera
Marañón Valley
O. cleefii
O. ecuadorensis
O. goeppingeri
O. obtusangulus
O. venezuelensis
500 km
A 1 Chirripo
B 2 Cocuy
B 5 Tota-Bijagual
C 3 La Rusia
C 4 Iguaque
C 6 Rabanal
C 7 Guerrero
D 8 Chingaza
D 9 Sumapaz
E 10 Purace
F 11 Azufral
F 12 Volcan Chiles
F 13 Mirador
F 14 Cotocachi
G 15 Cotopaxi
G 16 Llanganati
G 17 Alao Huamboya
H 18 Paramo de Las Cajas
H 19 Cuenca Limon
H 20 Cuenca Loja
I 21 Podocarpus
J 22 Cajamarca
J 23 Huascaran
K 24 Valdivia
L 25 Fiordo Peel
L 27 Morro Philippi
L 28 Magallanes
M 29 Tierra del Fuego
M 30 Isla de los Estados
M 31 Isla Navarino
M 32 Cabo de Hornos
N 26 Malvinas

Figure 2

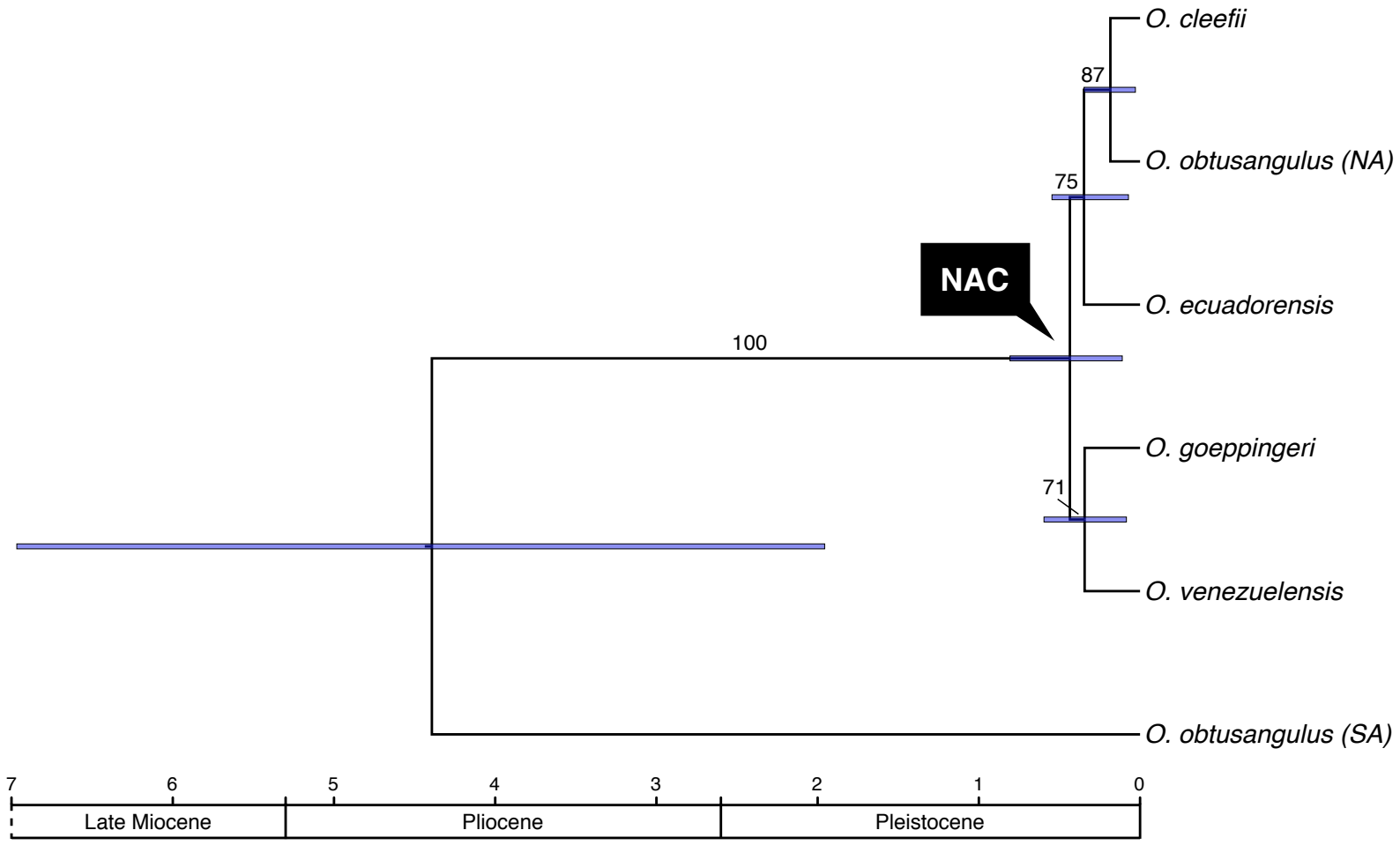

O. cleefii
87
O. obtusangulus (NA)
75
NAC
O. ecuadorensis
100
O. goeppingeri
71
O. venezuelensis
O. obtusangulus (SA)
7
6
5
4
3
2
1
0
Late Miocene
Pliocene
Pleistocene

Figure 3

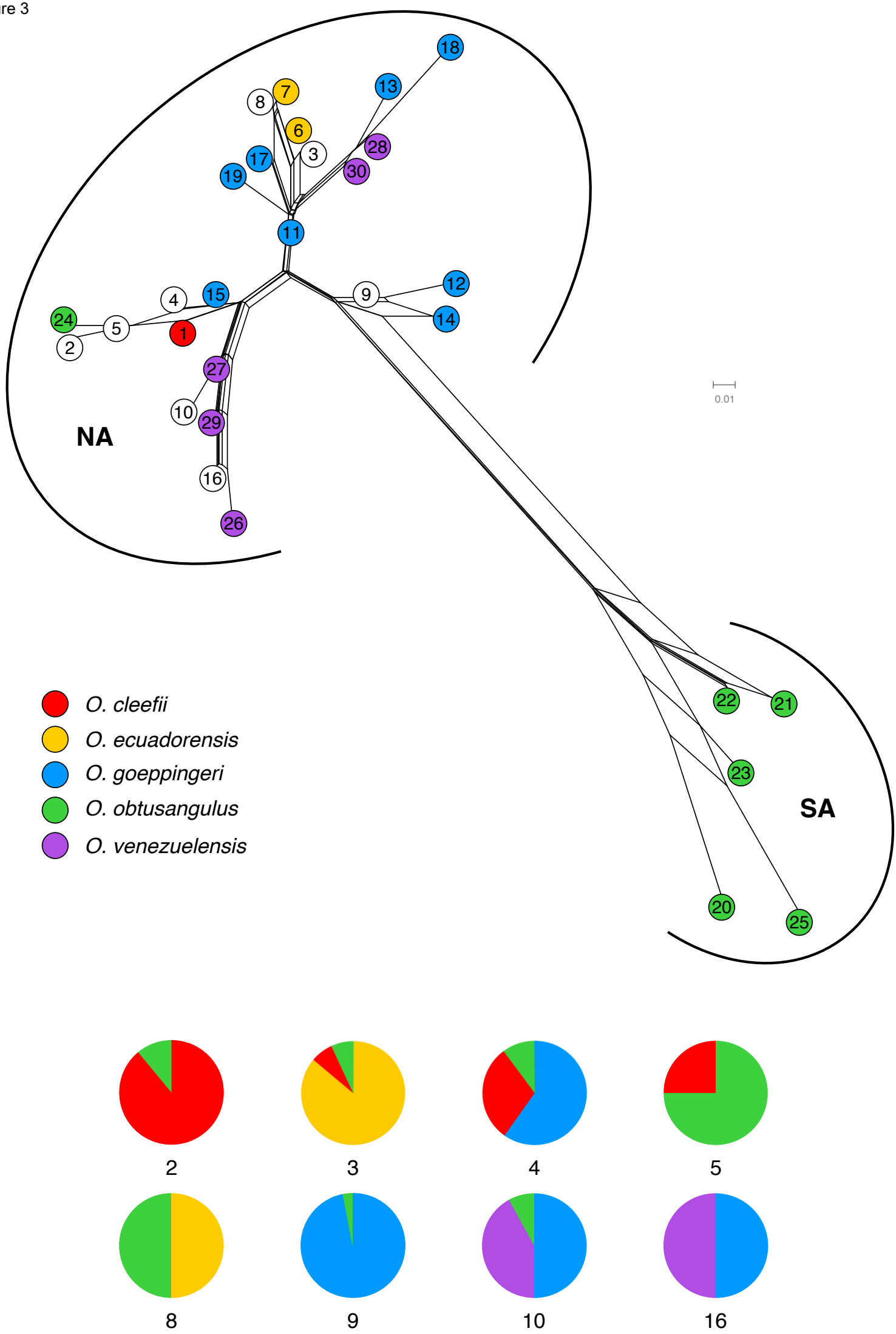
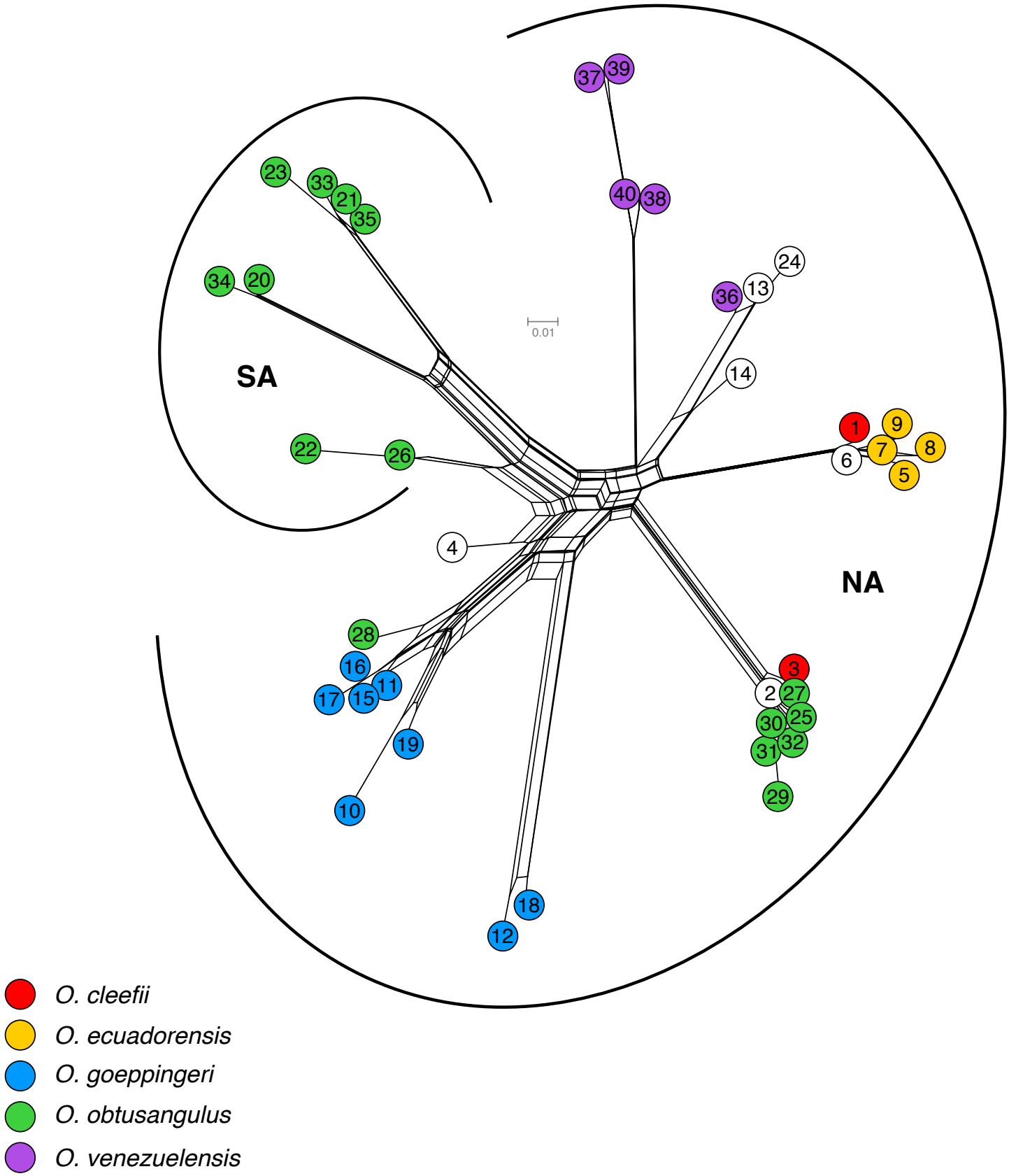

18
7
13
8
6
28
3
17
30
19
11
12
9
4
15
14
24
5
1
2
27
10
29
16
26
NA
0.01
22
21
23
SA
20
25
O. cleefii
O. ecuadorensis
O. goeppingeri
O. obtusangulus
O. venezuelensis
2
3
4
5
8
9
10
16

Figure 4



O. cleefii
O. ecuadorensis
O. goeppingeri
O. obtusangulus
O. venezuelensis

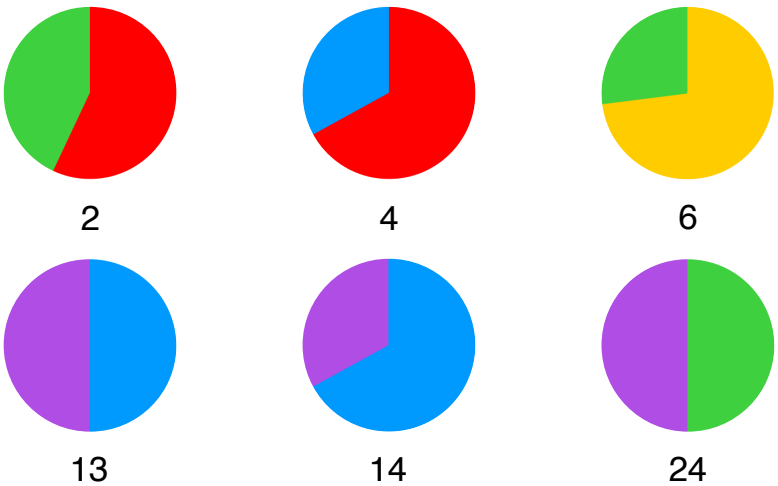


Figure 5